Mixture Dynamics and Regime Switching Diffusions with Application to Option Pricing

Alessandro Ramponi

Abstract In this paper we present a class of regime switching diffusion models described by a pair $(X(t), Y(t)) \in \mathbb{R}^n \times \mathcal{S}$, $\mathcal{S} = \{1, 2, \ldots, N\}$, $Y(t)$ being a Markov chain, for which the marginal probability of the diffusive component $X(t)$ is a given mixture. Our main motivation is to extend to a multivariate setting the class of mixture models proposed by Brigo and Mercurio in a series of papers. Furthermore, a simple algorithm is available for simulating paths through a thinning mechanism. The application to option pricing is considered by proposing a mixture version for the Margrabe Option formula and the Heston stochastic volatility formula for a plain vanilla.

Keywords Regime switching models · Mixture dynamics ·
Monte Carlo simulation · Option pricing

AMS 2000 Subject Classifications 60J60 · 60J27 · 65C05 · 60H30

1 Introduction

The need of considering price dynamics alternative to the classical Black-Scholes model for derivatives pricing is widely known. The stochastic variability of market parameters and in particular the empirical evidence of non constant surfaces of implied volatility in real markets require more realistic models for the assets dynamics. In the past years many new models have been proposed, each aiming to relax some of the restrictive assumption implied by the log-normal dynamic of the Black–Scholes model. These modifications typically consider explicit models for the local volatility, the addition of jump components and/or the introduction of stochastic volatility in the underlying diffusive dynamic.

A. Ramponi (✉)
Dept. S.E.F.eME.Q., University of Rome—Tor Vergata, via Columbia, 2, 00133 Rome, Italy
e-mail: ramponi@economia.uniroma2.it

In a series of papers Brigo and Mercurio (2000, 2002) and Brigo et al. (2003) proposed a class of one-dimensional diffusion models characterized by a local volatility function which implies a mixture distribution for the asset price. In particular, the case of log-normal mixtures has been considered and applied to various pricing problems. The main advantages of their model are i) that analytical formulas for European derivatives are readily available together with the corresponding Greeks and ii) an improved flexibility for the calibration to real market data. Furthermore, the explicit analytical form of the local volatility makes available Monte Carlo simulation to price path-dependent derivatives, through simple discretization schemes.

Our main contribution is to develop a mixture dynamic in a regime-switching diffusion framework in order to extend the class of mixture models proposed by Brigo and Mercurio. In fact, the main feature of our approach is twofold: it can be easily applied to multivariate processes and for important models sample paths can be exactly simulated. Furthermore, the (marginal) sample paths of our process may exhibit abrupt changes in their behavior nevertheless maintaining a mixture distribution for the diffusive component.

Regime switching process were introduced by Hamilton (1989, 1990) in a financial econometric context. The main idea consists in introducing a discrete and in general unobservable Markov chain which generates switches among a finite set of ``regimes´´: each regime characterizes a particular parameter set for the dynamic model. Indeed, empirical studies demonstrated that markets may randomly switch between low-volatility/high growth and/or high-volatility/low growth regimes. Correspondingly, there is a practical interest to develop methods for pricing and hedging options for regime switching underlying models. There has been a considerable progress in the case of standard European and/or American style options: see e.g Naik (1993), Di Masi et al. (1994), Bollen (1998), Guo (2001), Hardy (2001), Duan et al. (2002), Buffington and Elliott (2002), Guo and Zhang (2004), Liu et al. (2006), Yao et al. (2006), Jobert and Rogers (2006). Comparatively few results are available for exotic options: see Boyle and Draviam (2007) and the recent Elliott et al. (2007) who considered the pricing of volatility swaps in a regime-switching version of the Heston stochastic volatility model. Finally, switching Lévy processes have been considered in Elliott and Osakwe (2006).

In many of these applications the Markov chain is independent from the diffusive component and its generator is characterized by constant elements. In this paper we consider instead a state dependent regime-switching framework to get an explicit mixture dynamic. Mathematically the model we consider is described by the pair $(X(t), Y(t)) \in \mathbb{R}^n \times \mathcal{S}$, $\mathcal{S} = \{1, 2, \ldots, N\}$. The first component represents a "controlled diffusion" process since its drift and diffusion coefficients depend on the state of the second discrete component $Y(t)$. On the other hand $Y(t)$ is a "controlled Markov Chain" with finite state space and a generator depending on the continuous component $X(t)$. These kind of processes appear naturally in many engineering applications where the system under study may show abrupt changes in its structure and parameters (see e.g. Ghosh et al. 1993, 1997). Theoretical properties of these dynamical systems have been recently studied in Khasminskii et al. (2007). Our main result is a characterization of the Markov chain generator to get a target mixture marginal distribution for the diffusive component $X(t)$. This choice provides a class of models for which expected values can be evaluated as convex combinations of the expectations with respect to the components of the mixture. On the other hand,

the state dependent generator makes the regime-switching diffusion model difficult to work with computationally. Therefore a simulation scheme recently proposed for a general jump-diffusion process by Glasserman and Merener (2004) and based on a state dependent thinning mechanism has been adapted to our model. Numerical solution of switching (jump-)diffusions has been largely considered for constant Markov chain generators (see e.g. Yin et al. 2005; Mao et al. 2007).

The paper is organized as follows. In Section 2 the class of Regime Switching Diffusion models is introduced and the corresponding Kolmogorov Forward (or Fokker Planck) equation is used to obtain the target mixture probability. In Section 3 the problem of pricing a European plain vanilla option under our regime switching diffusion is presented. A simulation scheme for sampling paths is proposed together with illustrative examples involving log-normal mixtures. As a concrete application to option pricing models we extend two well-known formulas to our setting: the first one consider the Margrabe (or Exchange) Option and the second a plain vanilla under the Heston stochastic volatility model. Section 4 concludes the paper.

2 The Model and the Main Result

Let $(X(t), Y(t))$ be a two-component Markov process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with the natural filtration, such that the continuous component $X(t)$ takes values in $\mathbb{R}^n$ and its evolution is governed by the equation

$$dX(t) = \mu(t, X(t), Y(t))dt + \sigma(t, X(t), Y(t))dW(t) \tag{1}$$

where $W(t)$ is a d-dimensional brownian motion, $\mu : \mathbb{R}_+ \times \mathbb{R}^n \times \mathcal{S} \to \mathbb{R}^n$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \times \mathcal{S} \to \mathbb{R}^{n \times d}$, with $\sigma(t, x, i)\sigma(t, x, i)^T = a(t, x, i)$ and $Y(t)$ is a jump process with values in a finite state space $\mathcal{S}$ that we model as a marked point process (MPP).

We briefly recall here (see e.g. Runggaldier 2003) that a MPP can be characterized as a random measure $m(dz, dt)$ for which

$$\int_0^t \int_E H(z, s)m(dz, ds) = \sum_{n=1}^{N_t} H(Z_n, T_n)$$

where $\{T_n\}$ is a univariate point process on $\mathbb{R}^+$ and $\{Z_n\}$ is a sequence of random variables on a given measurable space $(E, \mathcal{E})$, N_t being the counting process. The corresponding stochastic intensity for m is a measure-valued process $\upsilon(dz, dt)$ for which

$$\int_0^t \int_E H(z, s)(m(dz, ds) - \upsilon(dz, ds))$$

is a martingale for each predictable process H. This setting has been introduced in the financial literature by Bijork et al. (1997).

We can now specify our jump process. Let $E = \{(i, j) \in \mathcal{S} \times \mathcal{S} : i \neq j\}$, $\mathcal{E} = 2^E$ be the finite mark space $(E, \mathcal{E})$ and consider the MPP $\nu(dt, dx)$ with the following stochastic intensity:

$$\upsilon(dt, dz) = \lambda(t, X(t), Y(t), dz)dt = \sum_{i \neq j} q_{ij}(t, X(t))1_{(Y(t-)=i)}\delta_{(i,j)}(dz)dt, \tag{2}$$

$\delta_{(i,j)}(dz)$ being the Dirac measure. The functions $q_{ij}(t, x)$ are positive and such that $\sum_{j=1, j\neq i}^{N} q_{ij}(t, x) = -q_{ii}(t, x)$, for $i = 1, \ldots, N$ and all t and x. In other words, we are considering a continuous-time and finite space Markov Chain (MC). The functions q_{ij} represent the elements of the intensity matrix of the MC: notice that we are allowing for a time and state dependent intensity. By defining for $z \in E$ the function $\varepsilon(z) = \varepsilon((i, j)) = j - i$, the dynamic of $Y(t)$ can be represented as

$$dY(t) = \int_E \varepsilon(z)\nu(dt, dz).$$

meaning that, if τ is a jump time and $Y(\tau-) = i$, then $Y(\tau) = Y(\tau-) + (j - i) = j$ with probability $q_{ij}(\tau, X(\tau))$.

Using the above notation, the dynamic of the $\mathbb{R}^n \times \mathcal{S}$-valued process $(X(t), Y(t))$ is given by the following SDE

$$dX(t) = \mu(t, X(t), Y(t))dt + \sigma(t, X(t), Y(t))dW(t)$$
$$dY(t) = \int_E \varepsilon(z)\nu(dt, dz) \tag{3}$$

where $W(t)$ is a standard d-dimensional Wiener process and $\nu(dt, dx)$ is the MPP on $(E, \mathcal{E})$ having compensator (2). We call Eq. 3 a Regime-Switching (RS) diffusion.

From now on we assume that the coefficients of our model satisfy the usual regularity conditions to ensure for any $t > 0$ the existence of 1) a strong solution and 2) a probability density function (pdf) for $(X(t), Y(t))$ on $\mathbb{R}^n \times \mathcal{S}$ w.r.t. the standard reference measure (see Skorohod 1989 for details). Furthermore we assume the following initial condition at $t = 0$: $Y(0) = Y_0$, where Y_0 is a random variable on $\mathcal{S}$ having probability mass function $\lambda_k = \mathbb{P}\{Y_0 = k\}$, $k \in \mathcal{S}$ and $X(0) = z \in \mathbb{R}^n$, in such a way

$$\mathbb{P}(X(t) \in A, Y(t) = i) = \sum_{k=1}^{N} \lambda_k \int_A p_{0,z,k}(t, x, i)dx \equiv \int_A \rho_0(t, x, i)dx,$$

for all $A \in \mathcal{B}(\mathbb{R})$ and $i \in \mathcal{S}$ where $\rho_0(t, x, i) = \sum_{k=1}^{N} \lambda_k p_{0,z,k}(t, x, i)$. Consequently, due to the Markov property, we have for a suitable integrable function $\Phi : \mathbb{R}^n \times \mathcal{S} \longrightarrow \mathbb{R}$

$$\mathbb{E}_0[\Phi(X(t), Y(t))] = \sum_{i=1}^{N} \int_{\mathbb{R}^n} \Phi(x, i)\rho_0(t, x, i)dx.$$

For any function $F(\cdot, \cdot, i) \in \mathcal{C}^{1,2}$, the generator of the process is

$$\mathcal{L}_t F(t, x, i) = \frac{\partial}{\partial t}F(t, x, i) + \frac{1}{2}\sum_{j,h=1}^{n} a_{jh}(t, x, i)\frac{\partial^2}{\partial x_j \partial x_h}F(t, x, i)$$

$$+ \sum_{j=1}^{n} \mu_j(t, x, i)\frac{\partial}{\partial x_j}F(t, x, i) + \sum_{j\neq i} q_{ij}(t, x)(F(t, x, j) - F(t, x, i)).$$

$$\tag{4}$$

Notice that $\sum_{j\neq i} q_{ij}(t, x)(F(t, x, j) - F(t, x, i)) = \sum_{j=1}^{N} q_{ij}(t, x)F(t, x, j)$. Standard stochastic calculus (see the Appendix for a sketch of the derivation) gives the Kolmogorov forward equation (KFE) for the density $\rho_0(t, x, i)$:

$$\frac{\partial}{\partial t}\rho_0(t, x, i) = \frac{1}{2}\sum_{j,h=1}^{n} \frac{\partial^2}{\partial x_j \partial x_h}(a_{jh}(t, x, i)\rho_0(t, x, i))$$

$$- \sum_{j=1}^{n} \frac{\partial}{\partial x_j}(\mu_j(t, x, i)\rho_0(t, x, i)) + \sum_{j=1}^{N} q_{ji}(t, x)\rho_0(t, x, j). \quad (5)$$

We now prove the following

Theorem 2.1 *For any $i \in S$ let $f_{(0,z)}^{(i)}(t, x)$ be density functions satisfying*

$$\frac{\partial}{\partial t} f_{(0,z)}^{(i)}(t, x) = \frac{1}{2}\sum_{j,h=1}^{n} \frac{\partial^2}{\partial x_j \partial x_h}(a_{jh}(t, x, i) f_{(0,z)}^{(i)}(t, x)) - \sum_{j=1}^{n} \frac{\partial}{\partial x_j}(\mu_j(t, x, i) f_{(0,z)}^{(i)}(t, x))$$

$$(6)$$

$f_{(0,z)}^{(i)}(0, x) = \delta_z(x), x \in \mathbb{R}^n$ *and let*

$$q_{ij}(t, x) = \alpha f_{(0,z)}^{(j)}(t, x)\lambda_j, \quad j \neq i, \quad q_{ii}(t, x) = -\sum_{j\neq i} q_{ij}(t, x), \quad (7)$$

$\alpha > 0$ *being an arbitrary constant. Then*

(i) $p_0(t, x, i) = \lambda_i f_{(0,z)}^{(i)}(t, x)$ *is the joint pdf of $(X(t), Y(t))$;*
(ii) *for any integrable function $\Phi : \mathbb{R}^n \times S \longrightarrow \mathbb{R}$ we have*

$$\mathbb{E}_0[\Phi(X(t), Y(t))] = \sum_{i=1}^{N} \lambda_i \mathbb{E}_0^{(i)}[\Phi(X(t), i)].$$

Proof The joint density of the solution $(X(t), Y(t))$ of Eq. 3 must satisfies the Kolmogorov Forward Eq. 5. By inserting the pdf $p_0(t, x, i)$ in Eq. 5 we have

$$\lambda_i \left(\frac{\partial}{\partial t} f_{(0,z)}^{(i)}(t, x) - \frac{1}{2}\sum_{j,h=1}^{n} \frac{\partial^2}{\partial x_j \partial x_h}(a_{jh}t, (x, i) f_{(0,z)}^{(i)}(t, x)) \right.$$

$$\left. + \sum_{j=1}^{n} \frac{\partial}{\partial x_j}(\mu_j(t, x, i) f_{(0,z)}^{(i)}(t, x)) \right) - \sum_{j=1}^{N} q_{ji}(t, x) f_{(0,z)}^{(j)}(t, x)\lambda_j = 0,$$

which clearly follows from the condition (6) and since from Eq. 7 we have for any i

$$\sum_{j=1}^{N} q_{ji}(t, x) f_{(0,z)}^{(j)}(t, x)\lambda_j = \sum_{j=1, j\neq i}^{N} \alpha f_{(0,z)}^{(i)}(t, x)\lambda_i f_{(0,z)}^{(j)}(t, x)\lambda_j + q_{ii} f_{(0,z)}^{(i)}(t, x)\lambda_i$$

$$= \alpha f_{(0,z)}^{(i)}(t, x)\lambda_i \sum_{j=1, j\neq i}^{N} f_{(0,z)}^{(j)}(t, x)\lambda_j$$

$$- f_{(0,z)}^{(i)}(t, x)\lambda_i \sum_{j=1, j\neq i}^{N} \alpha f_{(0,z)}^{(j)}(t, x)\lambda_j = 0.$$

Furthermore

$$\mathbb{E}_0[\Phi(X(t), Y(t)] = \sum_{i=1}^{N} \int_{\mathbb{R}^n} \Phi(x, i) p_0(t, x, i) dx = \sum_{i=1}^{N} \lambda_i \int_{\mathbb{R}^n} \Phi(x, i) f_{(0,z)}^{(i)}(t, x) dx \quad (8)$$

from which (ii) follows. □

Let us notice that the proof is unchanged for a non-constant multiplicative factor $\alpha \equiv \alpha(t, x)$.

By summing over i in the Eq. 5 we get

$$\frac{\partial}{\partial t} \sum_{i=1}^{N} \rho_0(t, x, i) = \frac{1}{2} \sum_{j,h=1}^{N} \frac{\partial^2}{\partial x_j \partial x_h} \sum_{i=1}^{N} a_{jh}(t, x, i) \rho_0(t, x, i)$$

$$- \sum_{j=1}^{N} \frac{\partial}{\partial x_j} \sum_{i=1}^{N} \mu_j(t, x, i) \rho_0(t, x, i), \quad (9)$$

since $\sum_{i=1}^{N} \sum_{j=1}^{N} q_{ji}(t, x) \rho_0(t, x, j) = 0$. On the other hand, for any SDE with drift $\tilde{\mu}_j(t, x)$ and diffusion coefficient $\tilde{a}_{jk}(t, x)$, $j, k = 1, \ldots, N$, the corresponding pdf $\tilde{\rho}_0(t, x)$ must satisfy

$$\frac{\partial}{\partial t} \tilde{\rho}_0(t, x) = \frac{1}{2} \sum_{j,h=1}^{n} \frac{\partial^2}{\partial x_j \partial x_h} (\tilde{a}_{jh}(t, x) \tilde{\rho}_0(t, x)) - \sum_{j=1}^{n} \frac{\partial}{\partial x_j} (\tilde{\mu}_j(t, x) \tilde{\rho}_0(t, x)). \quad (10)$$

Therefore, by comparing Eqs. 9 and 10, if

$$\tilde{a}_{jh}((t, x) = \sum_{i=1}^{N} \frac{a_{jh}(t, x, i) \rho_0(t, x, i)}{\sum_{\ell=1}^{N} \rho_0(t, x, \ell)}$$

$$\tilde{\mu}_j(t, x) = \sum_{i=1}^{N} \frac{\mu_j(t, x, i) \rho_0(t, x, i)}{\sum_{\ell=1}^{N} \rho_0(t, x, \ell)}$$

and provided $\tilde{\rho}_0(t, x) = \sum_{i=1}^{N} \rho_0(t, x, i)$ we get a mixture pdf for the solution $\mathcal{X}(t)$ of the SDE

$$d\mathcal{X}(t) = \tilde{\mu}(t, \mathcal{X}(t)) dt + \tilde{\sigma}(t, \mathcal{X}(t)) d\mathcal{W}(t),$$

$\tilde{\sigma}(t, x) \tilde{\sigma}^T(t, x) = \tilde{a}(t, x)$.

Now let $d = n = 1$. Under the hypothesis of Theorem 2.1 we get the mixture $p_m(t, x) = \sum_{i=1}^{N} \lambda_i f_{(0,z)}^{(i)}(t, x)$ as the pdf of the solution of the SDE

$$d\mathcal{X}(t) = \sum_{i=1}^{N} \frac{\mu(t, \mathcal{X}(t), i) \lambda_i f_{(0,z)}^{(i)}(t, \mathcal{X}(t))}{\sum_{\ell=1}^{N} \lambda_\ell f_{(0,z)}^{(\ell)}(t, \mathcal{X}(t))} dt + \sqrt{\sum_{i=1}^{N} \frac{\sigma^2(t, \mathcal{X}(t), i) \lambda_i f_{(0,z)}^{(i)}(t, \mathcal{X}(t))}{\sum_{\ell=1}^{N} \lambda_\ell f_{(0,z)}^{(\ell)}(t, \mathcal{X}(t))}} d\mathcal{W}(t)$$

$$(11)$$

as in Brigo and Mercurio (2000, 2002) (with $\mu(t, x, i) \equiv \mu$—see also Brigo et al. 2003). Clearly the conditions for the existence and uniqueness of the solution of the SDE (11) depend on the densities $f^{(i)}$'s.

It is worth noting that even if the (marginal) probability of the two continuous processes is the same mixture, their pathwise properties are very different.

3 Application to Option Pricing Models

The main advantages of considering convex linear combinations of densities as models for asset prices are easily summarized (see Brigo and Mercurio 2000): 1) analytical formulas for European derivatives are readily available for basic models together with the corresponding Greeks; 2) a potentially very large numbers of parameters to improve the flexibility for calibration to real market data. Furthermore, the explicit analytical form of the model dynamic makes available Monte Carlo simulation to price path-dependent derivatives, through simple discretization schemes.

The class of models introduced in the previous section, when considered as candidate price dynamics, according to Theorem 2.1 certainly shares these features, in a very general setting. Notice that the market we consider is incomplete in general and hence there is more than one pricing measure. Therefore in the following we assume that Eq. 3 characterize the dynamic of the factors driving the asset prices $S(t)$ under a given risk-neutral pricing measure Q, where $\nu(dt, dz)$ is the MPP on $(E, \mathcal{E})$ with intensity

$$\lambda(t, X(t), Y(t), dz) = \sum_{j \neq i} \alpha f^{(j)}(t, X(t))\lambda_j 1_{(Y(t-)=i)}\delta_{(i,j)}(dz) \tag{12}$$

and the $f^{(j)}(t, x)$ are the pdf of the solution of $dX(t) = \mu(t, X(t), j)dt + \sigma(t, X(t), j)dW(t)$, $\{\lambda_j\}_{j=1,...,N}$ being the probability mass function of Y_0. Here, along with asset prices the vector $X(t)$ can include also latent factors, as we will see in the last example, i.e. $X(t) = (S_1(t), \ldots, S_K(t), X_{K+1}(t), \ldots, X_n(t))$. As a consequence of the risk-neutrality assumption, we can restrict the drift term of the asset price component to be of the form $\mu(t, s, i) = \mu s$, where the parameter μ is specified by the pricing measure Q ($\mu = r - q$, where r is the risk-free rate and q the continuous dividend yield, $\mu = r - r_f$, r_f the foreign risk-free rate, for underlyings that are exchange rates—see Brigo et al. 2003). Finally, let $\Phi(s)$ be the integrable payoff of the derivative maturing at time T. Then, from Theorem 2.1, its price at time $t = 0$ is simply given by

$$c(0, T) = e^{-rT}\mathbb{E}_0[\Phi(S(T))] = \sum_{i=1}^{N} \lambda_i c_i(0, T), \tag{13}$$

where

$$c_i(0, T) = e^{-rT}\mathbb{E}_0^{(i)}[\Phi(S(T))]$$

is the time $t = 0$ price of the derivative according to the i-th state of the chain.

The case of one-dimensional log-normal mixture dynamics has been considered by Brigo and Mercurio (2000, 2002) and Brigo et al. (2003) and fully exploited for pricing European call/put options. In the following we present firstly a simulation scheme for RS diffusions to improve the classical discretization approach for generating sample paths from Eq. 11. Secondly, we propose two specific applications of the dynamic mixture model which permit to extend well-known option pricing formulas: the first

one consider the Margrabe (or Exchange) Option and the second extend the Heston stochastic volatility model to our setting.

3.1 Simulation

The state dependent intensity makes the RS diffusion model difficult to work with computationally. Glasserman and Merener (2004) proposed a class of discretization schemes for jump diffusion processes in which the jump component is modelled as a MPP with state-dependent intensity. Their main motivation relies on the fact that the change of measure technique for option pricing can introduce state dependent intensity when passing from objective to risk-neutral probability in a jump-diffusion model.

The main idea of their schemes is to enlarge the mark space and to "embed" the state-dependent jump component into a Poisson random measure on the enlarged mark space through a state dependent thinning mechanism. The Poisson random measure can be easily simulated and the thinning function accepts or rejects the generated marks with a probability depending on the state of the system.

To be definite, for our dynamic (3) characterized by the intensity (2), let $p(dz \times du, dt)$ be the Poisson random measure with mark space $E^* = E \times (0, 1)$ and intensity $v_P(z, u, t) = \lambda_0 h(z)$, $z \in E$, $u \in (0, 1)$: $h(z) = \sum_{z \in E} h_{ij} \delta_{(ij)}(dz)$ is a probability mass function on E ($h_{ij} > 0$, $\sum_{z \in E} h_{ij} = 1$) and λ_0 the arrival rate. The parameters h_{ij} and the rate λ_0 must be chosen in such a way

$$q_{ij}(t, x) < \lambda_0 h_{ij}, \quad (i, j) \in E. \tag{14}$$

for every $t > 0$ and $x \in \mathbb{R}^n$. The thinning function is then defined as

$$\theta(t, x, y, z, u) = \begin{cases} 1 & \text{if } u < \dfrac{q_{ij}(t, x) 1_{(y=i)}}{\lambda_0 h_{ij}} \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

As in Glasserman and Merener (2004), it can be proved that the process

$$d\tilde{Y}(t) = \int_E \int_0^1 \varepsilon(z) \theta(t, X(t-), Y(t-), z, u) p(dz \times du, dt) = \int_E \varepsilon(z) m(dz, dt),$$

where $m(dz, dt) = \int_0^1 \theta(t, x, y, z, u) p(dz \times du, dt)$ is equivalent to the one defined in Eq. 3.

The Euler simulation scheme (see e.g. Kloeden and Platen 1999), on the time interval $[0, T]$ for our RS dynamic proceed as follows. Let $\{\tau_k\}_k$ be a discrete set of times obtained as the superposition of the Poisson jump times and a deterministic grid on $[0, T]$. The value of the process $(X(\tau_k), Y(\tau_k))$ is iteratively computed using

$$X(\tau_{k+1}) = X(\tau_k) + \mu(X(\tau_k), Y(\tau_k)) \Delta \tau_k$$
$$+ \sigma(\tau_k, X(\tau_k), Y(\tau_k))(W(\tau_{k+1}) - W(\tau_k)) \tag{16}$$

$$Y(\tau_{k+1}) = Y(\tau_k) + \int_E \varepsilon(z) m(dz, dt) \tag{17}$$

The new state of the chain $Y(\tau_{k+1})$ is generated through the thinning mechanism only if τ_{k+1} is a point of the Poisson random measure, otherwise the jump term is

zero. For the RS mixture dynamic characterized by the intensity (12) this amounts to accept the switch from $Y(t-) = i$ to $Y(t) = j$ if

$$U < \frac{\alpha f_{(0,z)}^{(j)}(t, X(t))\lambda_j}{\lambda_0 h_{ij}} \tag{18}$$

where U is a uniform r.v. on $(0, 1)$. Clearly, other more sophisticated discretization schemes can be used for sampling the continuous component $X(t)$. We call the iterations (16), (17) and (18) the Regime Switching Thinning Algorithm (RSTA). In general, as proved in Glasserman and Merener (2004), under some conditions on the coefficients of the jump-diffusion equation and the set of test functions, the weak convergence order of this scheme is the same of the scheme used for the diffusive component of the system.

Remark 3.1 It is important to notice that if the continuous component $X(t)$ can be sampled exactly in each time interval $[\tau_k, \tau_{k+1}]$, the above scheme produces sample paths without discretization errors. This is a clear advantage compared to the purely diffusive representation (11) which requires approximation schemes for generating paths.

In the practical implementation of the RSTA we have the freedom of choosing the parameters α, λ_0 and the probabilities h_{ij}. While these last can always be selected to be the uniform law on E, i.e. $h_{ij} = 1/|E|$, in our experiments we set $\alpha \equiv \alpha(t, x) = (\sum_{k=1}^{N} \lambda_k f_{(0,z)}^{(k)}(t, x))^{-1}$ and $\lambda_0 = \max_{i,j}\{\lambda_j/h_{ij}\}$ in order to ensure condition (14). With such a choice of $\alpha(t, x)$, the RSTA resembles the Gibbs Sampler algorithm (see e.g. Robert and Casella 2004) since the switch from regime i to regime j at a jump time τ happens with probability (proportional to) the conditional probability of $Y(\tau)$ given $X(\tau)$.

As an illustrative example, let us consider the following dynamic:

$$dX(t) = X(t)\mu(Y(t)) + X(t)\sigma(Y(t))dW(t)$$

$$dY(t) = \int_E \varepsilon(z)\nu(dt, dz)$$

where $Y(t)$ is a two-state Markov chain, i.e. $E = \{(1, 2), (2, 1)\}$, $\mu_1, \mu_2 \in \mathbb{R}$ and $\sigma_1, \sigma_2 > 0$ are given parameters. For a given state $i = 1, 2$ of the chain, $X(t)$ is a GBM having a log-normal distribution

$$f_{(0,z)}^{(i)}(t, x) = \frac{1}{\sqrt{2\pi\sigma_i^2 tx}} \exp\left(-\frac{(\log(x) - m_i)^2}{2\sigma_i^2 t}\right), \quad x > 0$$

$$m_i = \log(z) + (\mu_i - \sigma_i^2/2)t.$$

According to Theorem 2.1 we consider the stochastic intensity (2) defined by the functions

$$q_{ij}(t, x) = \alpha(t, x) f_{(0,z)}^{(j)}(t, x)\lambda_j, \quad i, j = 1, 2, i \neq j, \quad \alpha(t, x) = \frac{1}{\sum_{k=1}^{2} \lambda_k f_{(0,z)}^{(k)}(t, x)}.$$

The law of Y_0 is $\mathbb{P}\{Y_0 = 1\} = 1 - \mathbb{P}\{Y_0 = 2\} = 2/5$ and $X(0) = 1$. Figure 1 shows the time evolution of the mixture pdf. In Fig. 2 the empirical densities obtained through the simulation scheme are compared with the pdf of $X(t)$ for various times
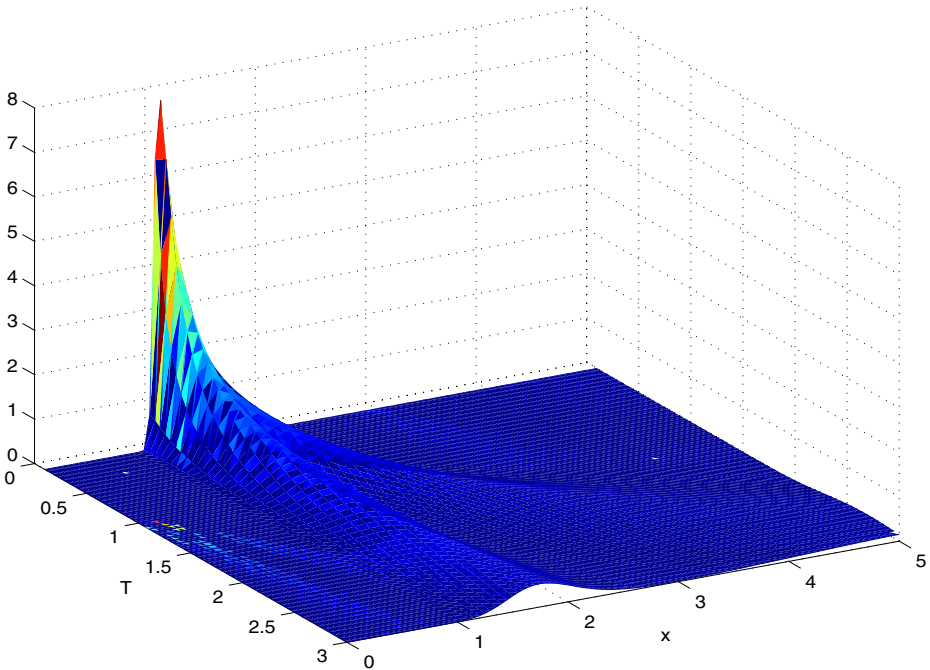
Fig. 1 Time evolution of the marginal pdf of $X(t)$. Here $z = 1$, $\mu_1 = .2$, $\mu_2 = .6$, $\sigma_1 = 0.1$, $\sigma_2 = .15$ and $T = 3$

in the interval $[0, T]$. The intensity of the Poisson random measure is $\lambda_0 = 1$ and the marks are sampled uniformly on E.

We test the efficiency of the our sampling algorithm by evaluating the price of a call option under a $N = 2$ component mixture of lognormals with parameters $\sigma = [.2 .5]$, $\lambda = [.5 .5]$ and $K = 100$, $T = 1$, $\mu_1 = \mu_2 \equiv r = .05$, $S_0 = 100$. The value of the call is $c = 16.1216$. We compare the bias and the RMSE of the estimates obtained with the RSTA to that obtained with the classical Euler-Maruyama dicretization algorithm for the Brigo and Mercurio SDE (EM-SDE) (a sample path of each model is depicted in Fig. 3):

$$dS(t) = S(t)\left(rdt + \sqrt{\sum_{i=1}^{2} \frac{\sigma_i \lambda_i \exp\left(-(\log(S(t)) - m_i)^2/(2\sigma_i^2 t)\right)}{\sum_{\ell=1}^{2} \lambda_\ell \exp\left(-(\log(S(t)) - m_\ell)^2/(2\sigma_\ell^2 t)\right)/\sigma_\ell}} dW(t)\right).$$

The algorithms were implemented in MatLab 6.5 (Release 13) by using the build-in generators for all the random variables involved.

We briefly recall that if $\hat{c}$ is an estimate of $c = \mathbb{E}[X]$, the bias is defined as $b = \mathbb{E}[\hat{c}] - c$ and the RMSE is the quantity $\sqrt{b^2 + \text{var}(\hat{c})}$. Furthermore, we consider the ratio of the computational times spent by the algorithms for generating all the paths, $r_{CT} = \text{CPUTIME(EM-SDE)/CPUTIME(RSTA)}$. The results reported in Table 1 appears encouraging.
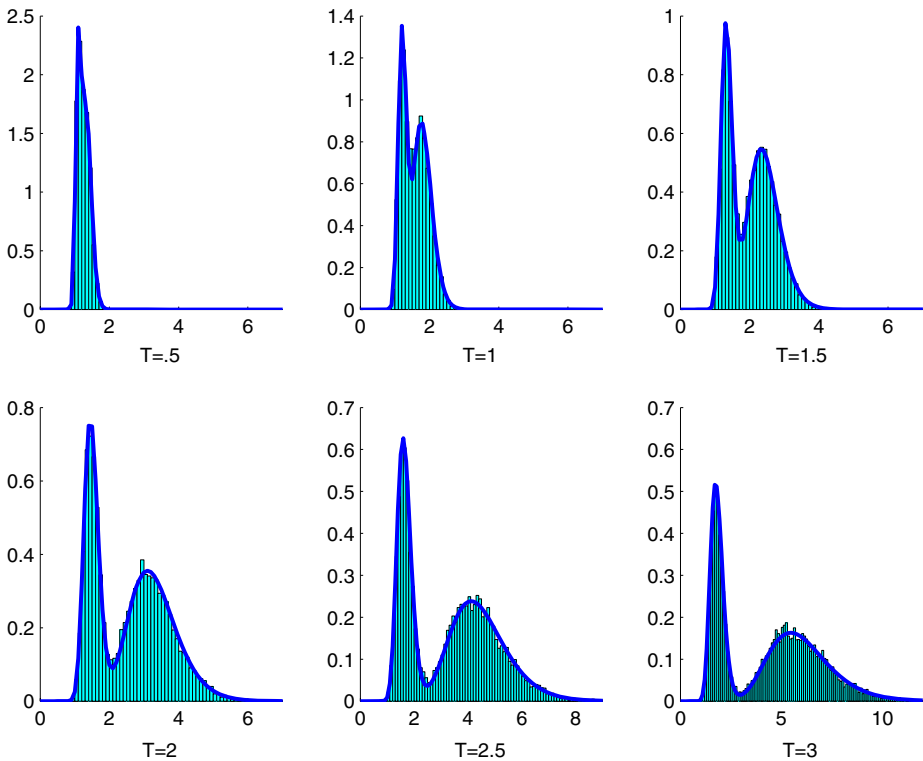
Fig. 2 Histograms of the sampled marginal process $X(t)$ compared with the marginal mixture $p_m(t, x) = \sum_{j=1}^{2} \lambda_j \frac{1}{\sqrt{2\pi\sigma_j^2 t x}} \exp\left(-\frac{(\log(x)-m_j)^2}{2\sigma_j^2 t}\right)$, $m_j = \log(x_0) + (\mu_j - \sigma_j^2/2)t$

3.2 Log-normal Mixtures for the Margrabe Option

As a first application of the results obtained in the previous section, we consider the following two-dimensional log-normal model

$$
\begin{aligned}
dS_t^1 &= S_t^1 \big(r dt + \sigma_1\big(\sqrt{1-\rho^2}dW_t^1 + \rho dW_t^2\big)\big) \\
dS_t^2 &= S_t^2 \big(r dt + \sigma_2 dW_t^2\big),
\end{aligned}
\tag{19}
$$

as the extension of the standard B&S asset price dynamics, under a risk-neutral measure. Here r is the risk-free rate, ρ is the correlation coefficient and (W_t^1, W_t^2) is a two-dimensional Brownian motion. The Exchange option gives the holder the right to exchange one asset for another: the corresponding payoff at the expiry time is $\Phi(S_T^1, S_T^2) = (S_T^1 - S_T^2)^+$. A pricing formula, under the previous dynamic, is obtained through standard arguments (see e.g. Musiela and Rutkowski 2005) giving

$$
c(t, T) = S_0^1 \mathcal{N}(d_1) - S_0^2 \mathcal{N}(d_2),
$$
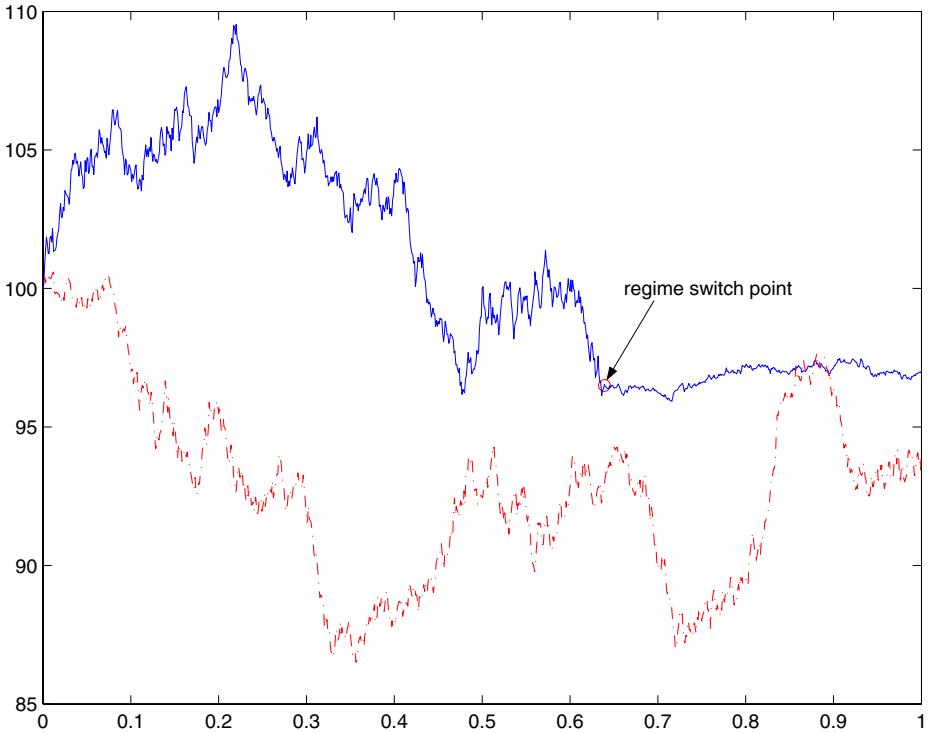
Fig. 3 Sample paths of the regime switching mixture model model (*blue line*) and the Brigo Mercurio model (*red line*)

known as Fisher–Margrabe Formula, where

$$d_{1,2} = \frac{\log S_0^1 / S_0^2 \pm \Sigma \cdot (T - t)/2}{\sqrt{\Sigma \cdot (T - t)}}, \quad \Sigma = \sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2.$$

Without loss of generality we can assume that $r = 0$. The formula strongly resembles the classical Black and Scholes formula where the correlation between assets gets incorporated in the volatility. Now, we can consider the Mixture Regime Switching dynamic by assuming stochastic volatilities $\sigma_1(Y(t))$, $\sigma_2(Y(t))$ and a stochastic corre-

Table 1 Estimated bias, (RMSE) and computational time r_{CT} for the Monte Carlo estimates of the call price. The time step length Δt for the Euler discretization scheme was chosen as $\Delta t = M^{-1/2}$

	50000	100000	150000	200000
RSTA	0.0987 (0.1721)	0.0533 (0.1130)	0.0350 (0.0886)	0.0128 (0.0714)
Euler	0.1652 (0.2180)	0.0821 (0.1081)	0.0238 (0.0914)	0.0169 (0.0827)
	$r_{CT} = 5.1991$	$r_{CT} = 4.0101$	$r_{CT} = 2.3539$	$r_{CT} = 2.0377$

lation $\rho(Y(t))$, taking values in the finite sets $\{\sigma_1(1), \ldots, \sigma_1(N)\}$, $\{\sigma_2(1), \ldots, \sigma_2(N)\}$ and $\{\rho(1), \ldots, \rho(N)\}$, respectively, as follows:

$$
\begin{aligned}
dS_t^1 &= S_t^1 \sigma_1(Y(t))\left(\sqrt{1 - \rho^2(Y(t))}dW_t^1 + \rho(Y(t))dW_t^2\right) \\
dS_t^2 &= S_t^2 \sigma_2(Y(t))dW_t^2 \\
dY(t) &= \int \varepsilon(z)\nu(dt, dz),
\end{aligned}
\tag{20}
$$

where $Y(t)$ has generator (12), the f^j's being two-dimensional log-normal densities. In particular, the bivariate r.v. $\begin{pmatrix} \log\left(S_t^1\right) \\ \log\left(S_t^2\right) \end{pmatrix}$ under regime i is normally distributed with expected value $\begin{pmatrix} \log\left(S_0^1\right) - \frac{\sigma_1^2(i)}{2}t \\ \log\left(S_0^2\right) - \frac{\sigma_2^2(i)}{2}t \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2(i) & \sigma_1(i)\sigma_2(i)\rho(i) \\ \sigma_1(i)\sigma_2(i)\rho(i) & \sigma_2^2(i) \end{pmatrix} t$. According to Theorem 2.1, if $Y(t)$ is the Markov chain with intensity (12), the resulting price is

$$
c_M(0, T) = \sum_{i=1}^N \lambda_i \mathbb{E}_0^{(i)}\left[\left(S_T^1 - S_T^2\right)^+\right] = \sum_{i=1}^N \lambda_i \left(S_0^1 \mathcal{N}(d_1(i)) - S_0^2 \mathcal{N}(d_2(i))\right)
\tag{21}
$$

where

$$
d_{1,2}(i) = \frac{\log S_0^1/S_0^2 \pm \Sigma(i) \cdot (T-t)/2}{\sqrt{\Sigma(i) \cdot (T-t)}}, \quad \Sigma(i) = \sigma_1^2(i) - 2\rho(i)\sigma_1(i)\sigma_2(i) + \sigma_2^2(i).
$$

It is worth noting that other kind of two-asset options, like the Chooser Option or some FX (quantos) options (see e.g. Musiela and Rutkowski 2005), can be treated in the same way.

3.3 Heston Stochastic Volatility Mixtures

Let us consider the Heston Stochastic Volatility (SV) model

$$
\begin{aligned}
dS(t) &= \mu S(t) + \sigma(t)S(t)dW^{(1)}(t) \\
dv(t) &= \kappa(\theta - v(t))dt + \gamma\sqrt{v(t)}\left(\rho dW^{(1)}(t) + \sqrt{1 - \rho^2}dW^{(2)}(t)\right), \quad \sigma^2(t) = v(t)
\end{aligned}
\tag{22}
$$

where $\left(W^{(1)}(t), W^{(2)}(t)\right)$ is a standard 2-dim Wiener process and $\rho \in [-1, 1]$ is their correlation coefficient.

The involved parameters are the initial volatility $\sigma(0) > 0$, the mean reversion rate $\kappa > 0$, the long run variance $\theta > 0$, the volatility of the variance (or volvol) $\gamma > 0$ and the correlation ρ. If $2\kappa\theta > \gamma^2$ (known as Feller condition), the variance process is always positive and cannot reach zero.

For this model the joint pdf of $(S(t), v(t))$ (in particular of the log-returns $R(t) = \log(S(t)/S(0))$ and $v(t)$) has been exactly derived by Dragulescu and Yakovenko (2002) by solving the corresponding Fokker–Planck equation, in terms of its Fourier

transform. We can consider the Mixture Regime Switching version of this model specified by the SDE

$$
\begin{aligned}
dS(t) &= rS(t) + \sigma(t)S(t)dW^{(1)}(t) \\
dv(t) &= \kappa(Y(t))(\theta(Y(t)) - v(t))dt + \gamma(Y(t))\sqrt{v(t)}(\rho(Y(t))dW^{(1)}(t) \\
&\quad + \sqrt{1 - \rho(Y(t))^2}dW^{(2)}(t)), \quad \sigma^2(t) = v(t) \\
dY(t) &= \int \varepsilon(z)\nu(dt, dz)
\end{aligned}
\tag{23}
$$

where $Y(t)$ is the Markov chain with intensity (12). The complete set of parameters for the model are therefore $\kappa(i)$, $\theta(i)$, $\gamma(i)$ (assumed to satisfy the conditions $2\kappa(i)\theta(i) > \gamma^2(i)$), $\rho(i)$ and λ_i, $i = 1\ldots, N$. According to Eq. 13 the value of an European call option with strike price K is simply

$$
c(0, T) = e^{-rT}\mathbb{E}_0\big[(S(T) - K)^+\big] = \sum_{i=1}^{N}\lambda_i e^{-rT}\mathbb{E}_0^{(i)}\big[(S(T) - K)^+\big] = \sum_{i=1}^{N}\lambda_i c_H^{(i)}(0, T)
\tag{24}
$$
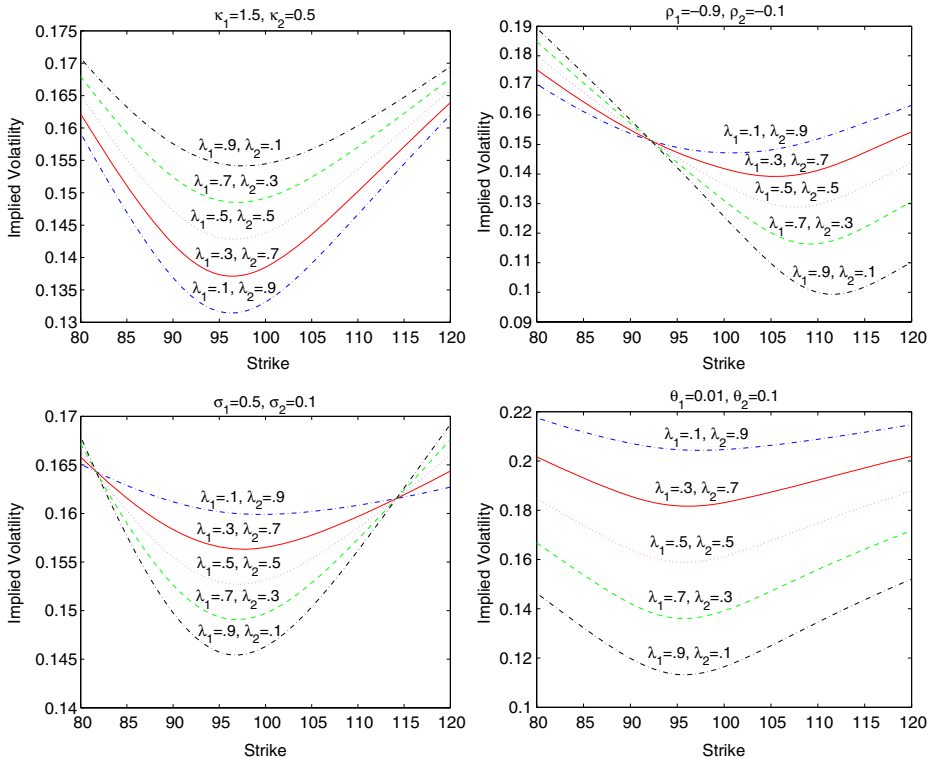
Fig. 4 Implied volatility curves for a mixture with two components and varying weights λ_1, λ_2

where (see Heston [1993] or Gatheral [2005])

$$c_H^{(i)}(0,\,T) = \frac{1}{2}\big(S_0 - e^{-rT}K\big) + \frac{1}{\pi}\int_0^{+\infty}\big(e^{rT}f_1(u) - Kf_2(u)\big)du,$$
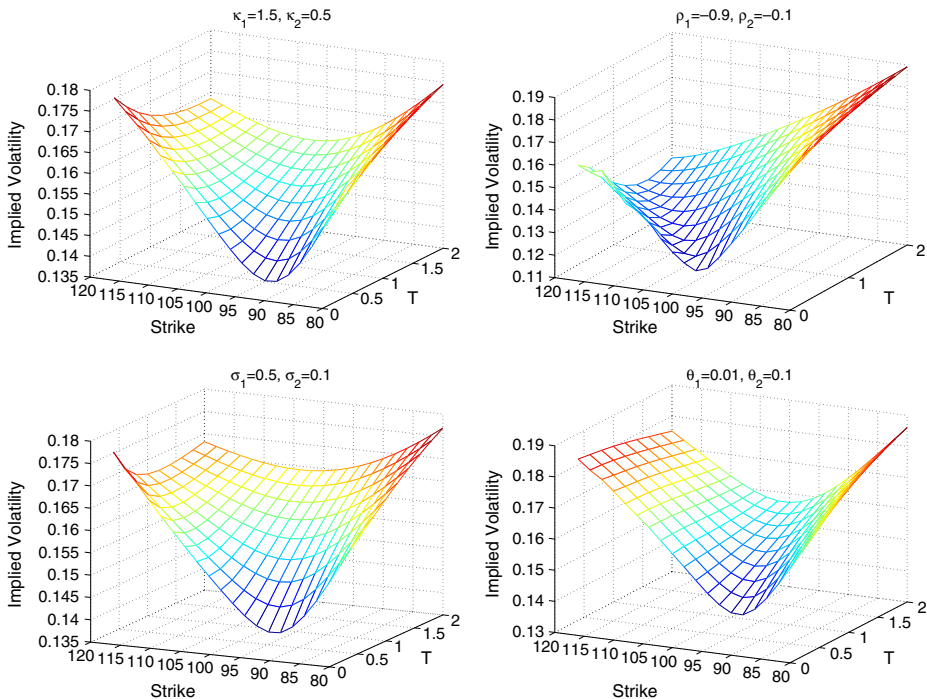
with

$$f_1^{(i)}(u) = \mathrm{Re}\left(\frac{\exp(-\mathrm{i}\log K)\phi^{(i)}(u - \mathrm{i},\,T)}{\mathrm{i}u\exp(rT)}\right), \quad f_2^{(i)}(u) = \mathrm{Re}\left(\frac{\exp(-\mathrm{i}\log K)\phi^{(i)}(u,\,T)}{\mathrm{i}u}\right),$$

i stands for the imaginary units and $\phi^{(i)}(u,\,T)$ is the characteristic function of the logarithm of the stock price process at time T for the i-th model given by
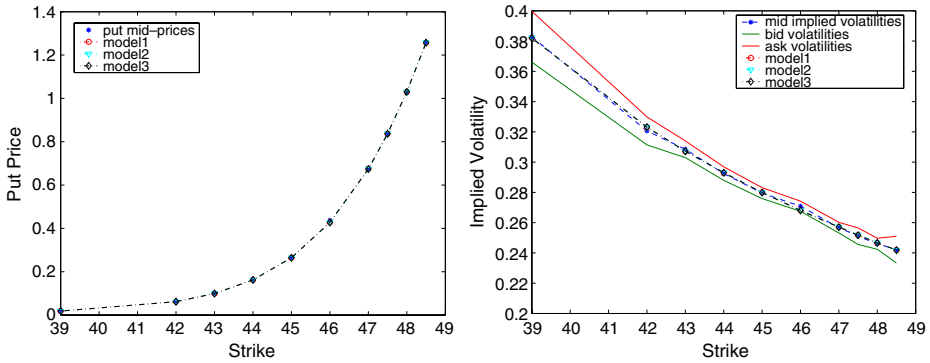
$$\begin{aligned}
\phi^{(i)}(u,\,T) = {}& \exp(\mathrm{i}u(\log S_0 + rT)) \\
& \times \exp(k(i)\theta(i)\gamma(i)^{-2} \\
& \times \left((k(i) - \rho(i)\gamma(i)u\mathrm{i} + d(i))T - 2\log\left(\frac{1 - g(i)e^{-d(i)T}}{1 - g(i)}\right)\right) \\
& \times \exp\left(\sigma_0^2\gamma(i)^{-2}(k(i) - \rho(i)\gamma(i)u\mathrm{i} + d(i))\frac{1 - e^{d(i)T}}{1 - g(i)e^{d(i)T}}\right),
\end{aligned}$$

$$d(i) = \sqrt{\rho(i)\gamma(i)u\mathrm{i} + \gamma(i)^2(u\mathrm{i} - u^2)}$$

$$g(i) = (k(i) - \gamma(i)\rho(i)u\mathrm{i} - d(i))/(k - \gamma(i)\rho(i)u\mathrm{i} + d(i)).$$

Fig. 5 Implied volatility surfaces for mixture with two components and weights $\lambda_1 = .6$, $\lambda_2 = .4$

Fig. 6 Calibration results for the mixture with two components. On the *left* we show the calibrated put prices vs. the observed ones, on the *right* the implied volatilities. The mean absolute errors for the three models Θ_1, Θ_2 and Θ_3 with respect to mid-prices and mid implied volatilities are (7.5845e−5, 1.1366e−5), (1.1579e−4, 9.9096e−6) and (9.7139e−5, 9.6943e−6) respectively

The impact of the mixture modeling over the option prices can be valued through the implied volatility curves. In Fig. 4 we generate prices according to the mixture of Heston SV models by varying only one parameter at time and fixing all the others in the convex combination. We consider mixtures with two components for different set of weights. The two values of the parameter which is varied in the mixture are reported on the top of each plot. The other fixed values are: $S(0) = 90$, $v(0) = 0.0175$, $r = 0.05$, $T = 1$ and $\kappa = 1.15$, $\theta = 0.04$ $\sigma = 0.39$, $\rho = −0.1$. As it can be seen, a large variety of shapes can be obtained. Similarly, in Fig. 5, implied volatility surfaces are plotted for a 2-components mixture with weights $\lambda_1 = .6$, $\lambda_2 = .4$.

As a simple example of calibration of the model in a real setting, we use the same market data as in Brigo and Mercurio (2000) (a set of put options bid - ask implied volatilities on the MIB30 equity index). We calibrate a $N = 2$ components mixture of Heston stochastic volatility models through a non-linear least square algorithm over the put mid-prices. We choose to calibrate three slightly different mixtures: the full model with parameters $\Theta_1 = (\kappa, \theta, \sigma_1, \sigma_2, \rho_1, \rho_2, \lambda, v_0^1, v_0^2)$ and two reduced ones with parameters $\Theta_2 = (\kappa, \theta, \sigma, \rho_1, \rho_2, \lambda, v_0^1, v_0^2)$ and $\Theta_3 = (\kappa, \theta, \rho, \sigma_1, \sigma_2, \lambda, v_0^1, v_0^2)$ respectively. The results obtained were very similar in terms of the mean absolute errors with respect to both the put mid-prices and the mid-implied volatilities (see Fig. 6). In particular we get $\hat{\Theta}_1 = 0.5587, 0.2822, 0.8999, −0.9998, 0.5789, −0.7776, 0.0323, 0.4521, 0.0504$) for the full model, $\hat{\Theta}_2 = (0.6188, 0.3239, 0.6065, −0.9998, −0.7371, 0.0302, 0.4999, 0.0493)$ and $\hat{\Theta}_3 = (0.5493, 0.2887, 0.8999, −0.7767, 0.5843, 0.0293, 0.4999, 0.0506)$ for the reduced ones.

4 Conclusions

In this paper we have presented a class of regime switching models described by a pair $(X(t), Y(t)) \in \mathbb{R}^n \times \mathcal{S}$, $\mathcal{S} = \{1, 2, \ldots, N\}$, $Y(t)$ being a controlled Markov chain, for which the marginal probability of the diffusive component $X(t)$ is a given mixture.

This regime switching framework permits to extend to a multidimensional setting the one-dimensional mixture dynamics proposed by Brigo and Mercurio in a series of papers. A simple simulation scheme for the trajectories of the process has been also proposed, based on the MPP representation of the Markov chain. As an application to the problem of pricing a European style option we have presented a log-normal mixture model for the (two-dimensional) Margrabe option and the mixture version of the Heston SV model. These models have never been considered in the literature, at least to the author's knowledge.

As a final comment for future research, we want to address two main points. Firstly, from a statistical perspective, the calibration of these mixtures, in particular the Heston SV mixture model, on real market data would be very interesting, both from a cross-sectional and a time series point of view (see e.g. Frühwirth-Schnatter 2006). Empirical studies comparing the performances of alternative models for option pricing are certainly of a great practical interest. More generally, the regime switching framework introduced is general enough and can be applied to other classes of financial problems. In particular, it would be interesting to explore the application of mixture models for the pricing of more complex derivatives, also in an interest rates setting, intensity based risk models, portfolio management and so on.

Appendix

Sketch of the derivation of the Kolmogorov Forward equation for the RS dynamic (3).

From an appropriate version of Ito's formula, we get for a suitable function F that

$$F(t, X(t), Y(t)) - \int_0^t (\mathcal{L}_s F)(s, X(s), Y(s)) ds$$

is a martingale. By taking the expectation and interchanging the order of integration we get

$$\mathbb{E}_0[F(t, X(t), Y(t))] = \mathbb{E}_0[F(0, X(0), Y_0)] + \int_0^t \mathbb{E}[(\mathcal{L}_s F)(s, X(s), Y(s))] ds$$

that is

$$\sum_{i=1}^N \int_{\mathbb{R}^n} F(t, x, i) \rho_0(t, x, i) dx = \sum_{k=1}^N \lambda_k F(0, z, k)$$

$$+ \int_0^t \left(\sum_{i=1}^N \int_{\mathbb{R}^n} (\mathcal{L}_s F)(s, x, i) \rho_0(t, x, i) dx \right) ds.$$

Let $\mathcal{L}^*$ be the differential operator satisfying

$$\sum_{i=1}^N \int_{\mathbb{R}^n} (\mathcal{L}_t f)(t, x, i) \rho_0(t, x, i) dx = \sum_{i=1}^N \int_{\mathbb{R}^n} f(t, x, i) (\mathcal{L}_t^* \rho_0)(t, x, i) dx \qquad (25)$$

for every suitable function f. Then we have

$$\sum_{i=1}^{N} \int_{\mathbb{R}^n} F(t, x, i)\rho_0(t, x, i)dx = \sum_{k=1}^{N} \lambda_k F(0, z, k)$$

$$+ \int_0^t \left(\sum_{i=1}^{N} \int_{\mathbb{R}^n} F(s, x, i)(\mathcal{L}_s^* \rho_0)(s, x, i)dx \right) ds$$

which is the weak form of the equation

$$\frac{\partial}{\partial t} \rho_0(t, x, i) = (\mathcal{L}_t^* \rho_0)(t, x, i) \tag{26}$$

usually named the Kolmogorov forward, or Fokker–Planck, equation. It remains to determine $\mathcal{L}^*$ from the condition (25). By defining

$$(\tilde{\mathcal{L}}_t f)(x, i) = \frac{\partial}{\partial t} f(t, x, i) + \frac{1}{2} \sum_{j,k=1}^{n} a_{jk}(t, x, i) \frac{\partial^2}{\partial x_j \partial x_k} f(t, x, i)$$

$$+ \sum_{j=1}^{n} \mu_j(t, x, i) \frac{\partial}{\partial x_j} f(t, x, i)$$

in such a way $(\mathcal{L}_t f)(x, i) = (\tilde{\mathcal{L}}_t f)(x, i) + \sum_{j=1}^{N} q_{ij}(t, x) f(t, x, j)$, we have

$$\sum_{i=1}^{N} \int_{\mathbb{R}^n} (\tilde{\mathcal{L}}_t f)(x, i)\rho_0(t, x, i)dx + \sum_{i=1}^{N} \int_{\mathbb{R}^n} \sum_{j=1}^{N} f(t, x, j)q_{ij}(t, x)\rho_0(t, x, i)dx$$

$$= \sum_{i=1}^{N} \int_{\mathbb{R}^n} f(t, x, i)(\tilde{\mathcal{L}}_t^* \rho_0)(t, x, i)dx + \sum_{j=1}^{N} \int_{\mathbb{R}^n} f(t, x, j) \sum_{i=1}^{N} q_{ij}(t, x)\rho_0(t, x, i)dx$$

$$= \sum_{i=1}^{N} \int_{\mathbb{R}^n} f(t, x, i)(\tilde{\mathcal{L}}_t^* \rho_0)(t, x, i)dx + \sum_{i=1}^{N} \int_{\mathbb{R}^n} f(t, x, i) \sum_{j=1}^{N} q_{ji}(t, x)\rho_0(t, x, j)dx$$

$$= \sum_{i=1}^{N} \int_{\mathbb{R}^n} f(t, x, i)(\mathcal{L}_t^* \rho_0)(t, x, i)dx$$

where

$$(\mathcal{L}_t^* \rho_0)(t, x, i) = (\tilde{\mathcal{L}}_t^* \rho_0)(t, x, i) + \sum_{j=1}^{N} q_{ji}(t, x)\rho_0(t, x, j)$$

and (see e.g. Gardiner 1983)

$$(\tilde{\mathcal{L}}_t^* \rho_0)(t, x, i) = \frac{1}{2} \sum_{j,h=1}^{n} \frac{\partial^2}{\partial x_j \partial x_h} (a_{jh}(x, i)\rho_0(t, x, i)) - \sum_{j=1}^{n} \frac{\partial}{\partial x_j} (\mu_j(x, i)\rho_0(t, x, i))$$

for every $i = 1, \ldots, N$. From Eq. 26, we finally get the Kolmogorov forward equation of our process:

$$\frac{\partial}{\partial t}\rho_0(t, x, i) = (\tilde{\mathcal{L}}_t^* \rho_0)(t, x, i) + \sum_{j=1}^{N} q_{ji}(t, x)\rho_0(t, x, j). \tag{27}$$

References

Bijork T, Kabanov Y, Runggaldier W (1997) Bond market in the presence of marked point processes. Math Financ 7(2):211–223

Bollen NPB (1998) Valuing options in regime-switching models. J Deriv 6:38–49

Boyle P, Draviam T (2007) Pricing exotic options under regime switching. Insur Math Econ 40: 267–282

Brigo D, Mercurio F (2000) A mixed-up smile. Risk 13:123–126

Brigo D, Mercurio F (2002) Lognormal-mixture dynamics and calibration to market volatility smiles. Int J Theor Appl Financ 5(4):427–446

Brigo D, Mercurio F, Sartorelli G (2003) Alternative asset-price dynamics and volatility smile. Quantitative Finance 3(3):173–183

Buffington J, Elliott RJ (2002) American options with regime switching. Int J Theor Appl Financ 5:497–514

Di Masi GB, Kabanov YM, Runggaldier WJ (1994) Mean-variance hedging of options on stocks with Markov volatility. Theory Probab Appl 39:173–181

Dragulescu AA, Yakovenko VM (2002) Probability distribution of returns in the Heston model with stochastic volatility. Quantitative Finance 2:443–453

Duan JC, Popova I, Ritchken P (2002) Option pricing under regime switching. Quantitative Finance 2:1–17

Elliott RJ, Osakwe CJU (2006) Option pricing for pure jump processes with Markov switching compensators. Finance Stoch 10:250–275

Elliott RJ, Siu TK, Chan L (2007) Pricing volatility swaps under Heston's stochastic volatility model with regime switching. Appl Math Financ 14(1):41–62

Frühwirth-Schnatter S (2006) Finite mixture and Markov switching models. Springer series in statistics. Springer, New York

Gardiner CW (1983) Handbook of stochastic methods. Springer, Berlin

Gatheral J (2005) The volatility surface: a practioner's guide. Wiley Finance, New York

Ghosh MK, Arapostathis A, Marcus SI (1993) Optimal control of switching diffusions with application to flexible manufacturing systems. SIAM J Control Optim 31:1183–1204

Ghosh MK, Arapostathis A, Marcus SI (1997) Ergodic control of switching diffusions. SIAM J Control Optim 35:1952–1988

Glasserman P, Merener N (2004) Convergence of a discretization scheme for jump-diffusion process with state-dependent probability. J R Stat Soc Ser 32(2):443–458

Guo X (2001) Information and option pricing. Quantitative Finance 1:38–44

Guo X, Zhang QZ (2004) Closed-form solutions for perpetual American put options with regime switching. SIAM J Appl Math 64:2034–2049

Hamilton JD (1989) A new approach to the economic analysis of non-stationary time series. Econometrica 57:357–384

Hamilton JD (1990) Analysis of time series subject to changes in regime. J Econom 45:39–70

Hardy MR (2001) A regime switching model of a long term stock-returns. N Am Actuar J 3:185–211

Heston SL (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financ Stud 6:327–343

Jobert A, Rogers LCG (2006) Option pricing with Markov-Modulated dynamics. SIAM J Control Optim 44(6):2063–2078

Khasminskii RZ, Zhu C, Yin G (2007) Stability of regime-switching diffusions. Stoch Process Their Appl 117:1037–1051

Kloeden PE, Platen E (1999) Numerical solution of stochastic differential equation. Springer, New York

Liu RH, Zhang Q, Yin G (2006) Option pricing in a regime-switching model using the fast Fourier transform. J Appl Math Stoch Anal 22:1–22

Mao X, Yuan C, Yin G (2007) Approximations of euler-maruyama type for stochastic differential equations with markovian switching, under non-lipschitz conditions. J Computat Appl Math 205:936–948

Musiela M, Rutkowski M (2005) Martingale methods in financial modelling. Springer, New York

Naik V (1993) Option valuation and hedging strategies with jumps in the volatility of asset returns. J Finance 48:1969–1984

Robert CP, Casella G (2004) Monte Carlo statistical methods. Springer, New York

Runggaldier WJ (2003) Jump-diffusion models. In: Rachev ST (ed) Handbook of heavy tailed distributions in finance. Handbooks in finance, book 1, (W.Ziemba Series Ed.). Elesevier/North-Holland, Amsterdam, pp 169–209

Skorohod AV (1989) Asymptotic methods in the theory of stochastic differential equations. American Mathematical Society, Providence

Yao DD, Zhang Q, Zhou XY (2006) A regime-switching model for European options. In: Yan HM, Yin G, Zhang Q (eds) Stochastic processes, optimization, and control theory applications in financial engineering, queueing networks, and manufacturing systems. Springer, New York, pp 281–300

Yin G, Song Q, Zhang Z (2005) Numerical solutions for jump-diffusions with regime switching. Stochastics 77:61–79